ENVIRONMETRICS
Environmetrics, **10**, 547–564 (1999)

CHANGE-POINT METHODS FOR WEIBULL MODELS WITH APPLICATIONS TO DETECTION OF TRENDS IN EXTREME TEMPERATURES

V. K. JANDHYALA[1]*, S. B. FOTOPOULOS[2] AND N. EVAGGELOPOULOS[3]

[1]*Department of Pure & Applied Mathematics and Program in Statistics, Washington State University, Pullman, WA 99164, USA*
[2]*Department of Management & Decision Sciences and Program in Statistics, Washington State University, Pullman, WA 99164, USA*
[3]*Department of Management & Decision Sciences, Washington State University, Pullman, WA 99164, USA*

SUMMARY

We develop change-point methodology for identifying dynamic trends in the scale and shape parameters of a Weibull distribution. The methodology includes asymptotics of the likelihood ratio statistic for detecting unknown changes in the parameters as well as asymptotics of the maximum likelihood estimate of the unknown change-point. The developed methodology is applied to detect dynamic changes in the minimum temperatures of Uppsala, Sweden. Copyright © 1999 John Wiley & Sons, Ltd.

KEY WORDS likelihood ratio statistic; maximum likelihood estimate; Weibull distribution; minimum temperatures

1. INTRODUCTION

In the area of environmental monitoring and assessment, environmentalists frequently encounter the problem of identifying dynamic trends (non-stationarities) in location, scale and/or shape of a proposed model. This has been recently emphasized by Jaruskova (1997). Katz and Brown (1992) earlier demonstrated how identification of changes in variability (scale and/or shape) help better predict the occurrence of extreme events in a changing climate. Hot spells and droughts are a frequent cause of adverse social impact to both humans and animals alike. For instance, Glantz (1987) studied the recurrent episodes of famine in Africa. Similarly, the primary impacts of climate on society result from occurrence of such extreme events. A hot spell during the summer of 1983 in the mid-western US resulted in a substantial decrease in corn yields (Mearns *et al.* 1984). Deep freezes during the winters of 1983 and 1985 killed a significant fraction of the citrus trees in the state of Florida (Miller and Glantz 1988). Accurate predictions of such extreme events will enable the society to be better prepared to cope with the consequences. Katz and Brown (1992) have shown conclusive evidence that the prediction of such extremities lies largely in identifying changes in the variability (scale and/or shape) of the climate conditions and to a lesser extent on scenarios of changes in the average.

* Correspondence to: V. K. Jandhyala, Department of Pure and Applied Mathematics and Program in Statistics, Washington State University, Pullman, WA 99164, USA.

CCC 1180–4009/99/050547–18$17·50

Received 5 January 1999
Accepted 14 March 1999

Jaruskova (1997) in a recent article discussed the role and importance of change-point methods for identifying dynamic trends (non-stationarities) in parameters such as mean and/or variability of environmental factors. Several studies are available in the literature where change-point methodology has been effectively implemented in identifying changes in the parameters of environmental models. Potter (1981) applied the methodology to detect changes in the mean level of 100-year precipitation series from the North-East US. MacNeill *et al.* (1991) examined the annual discharges of the Nile River at Aswan for changes in the mean level, as well as in the serial correlation structure, and found evidence for changes in both. This data has been earlier analyzed by many authors, including Cobb (1978). Recently, Brillinger (1994) analyzed the Nile River data again based on a wavelet approach. Brillinger (1994, 1997) also analyzed the data on monthly stages of the Rio Nigro River at Manus, Brazil from 1903 to 1992, and found little evidence for changes in the mean. Thus, change-point methodology has become an indispensable tool for scientists involved with modeling and monitoring environmental data.

Jaruskova (1994) and Gombay and Horvath (1997) analyzed the monthly averages of Nacetinsky Creek during the years 1951–1990. They have considered a lognormal model for the monthly averages data with changes in the scale and shape parameters over time. Since change-point methods for parameters of a lognormal model are not yet available in the literature, both Jaruskova (1994) and Gombay and Horvath (1997) carried out the change-point analysis by applying the log transformation and thus obtaining normality. Jaruskova (1994), however, pointed out that the log transformation leads to interpretation problems. For example, since both mean and variance of the normal distribution involve the scale as well as shape parameters of the lognormal distribution, change in mean only for the transformed data is not easily interpretable.

The Weibull distribution is also frequently applied for modeling and analyzing data in environmental sciences and elsewhere. It was developed in 1939 by the Swedish physicist Waloddi Weibull in order to describe the behavior of the breaking strength of materials. Specifically, Weibull (1939) demonstrated a statistical analysis that was particularly effective in modeling experimental fatigue data. Since his early pioneering work, the Weibull distribution has been used in various fields such as engineering, business, forestry, hydrology, biology and other related applied areas. In particular, the Weibull distribution has been successfully used in the probabilistic study of environmental factors. Essenwanger (1976), illustrated the applicability of several statistical distributions including the Weibull.

It is well known that extreme values of independent data intrinsically follow a Weibull distribution. As a result the Weibull distribution is widely applied to model data on climatological factors such as maximum/minimum daily temperatures. It also found immense applications for modeling cumulative damage due to fatigue in materials. In all these applications, one is naturally interested in incorporating dynamic effects on both scale and shape parameters of the Weibull distribution. Change-point methods that apply to the parameters of a Weibull model, however, have not yet been developed in the literature. The lack of such methods has been explicitly pointed out by de Rijk *et al.* (1990). While analyzing data on long term effects of dental materials caused by their exposure to an environment of oral fluids such as food simulating liquids, de Rijk *et al.* (1990) were interested in identifying changes in the growth rates of stress in dental specimens. They found it appropriate to model the stress growths by a two-parameter Weibull distribution. Thus, they were clearly looking for change-point methodology for identifying changes in the Weibull parameters. Pointing out that such methods were not available in the statistical literature, they carried out their analysis based on purely visual identifications.

The goal of this paper is to address this lack of availability and develop methods for both detection as well as estimation of unknown change-points in Weibull parameters. Adapting the likelihood principle, we first derive the null asymptotic distribution of the likelihood ratio statistic for the detection of an unknown change-point. We then obtain the asymptotic distribution of the maximum likelihood estimate (mle) of the change-point. The asymptotic distribution of the mle enables one to construct confidence interval estimates of any desired level. The parallel problem of developing change-point methodology for the lognormal distribution is currently in progress.

We apply the developed methodology to time series data on daily minimum temperatures at Uppsala, Sweden. Leadbetter *et al.* (1983) extracted this data from the original manuscripts compiled by Sverker Hellstrom. The methodology shows evidence of more than one change occurring in both the scale and shape parameters of a fitted Weibull model. Residual analysis validates independence among this data, which is a key assumption in our formulation. Realistically, one might wish to develop the methodology for identifying dynamic trends in time series data under serial correlations. While some recent advances have been made for the detection of change-points in serially correlated processes (Davis *et al.* 1995), the problem of deriving the asymptotic distribution the estimate of an unknown change-point under correlations remains open even for standard distributions such as normal and exponential distributions.

2. DETECTION OF CHANGES IN THE WEIBULL PARAMETERS

Several authors addressed the likelihood ratio method for detecting an unknown change-point in time ordered data. Hawkins (1977) derived the likelihood ratio statistic for detecting change in the mean of a sequence of normally distributed random variables and computed the finite sample exact null distribution of the statistic. Worsley (1986) extended the computation of the finite sample distribution to a sequence of exponential random variables. James *et al.* (1987) and Kim and Siegmund (1989) found some approximations for the null exceedence probabilities of the likelihood ratio statistic. Yao and Davis (1986) and Haccou *et al.* (1988) established the asymptotic distribution of the likelihood ratio statistic in the normal and exponential cases, respectively. They found the null asymptotic distribution to be of a double exponential type extreme value distribution. Gombay and Horvath (1990) extended this asymptotic result to the case of testing for the change-point in the mean of a sequence of independent random variables having a general distribution. Furthermore, the double exponential extreme value limiting distribution has been shown to hold more generally by Horvath (1993a, 1993b), Gombay and Horvath (1994a, 1994b, 1996a, 1996b) and Davis *et al.* (1995). These general situations include testing for a change-point in multiple parameters, in the parameters of a general linear regression model and also in the parameters of an auto-regressive process.

In this section, we establish (see Appendix for proof) the validity of the double exponential distribution for the log likelihood ratio statistic for detecting unknown changes in Weibull parameters. We begin by letting $Y_1, \ldots, Y_n$ be a sequence of time-ordered independent random variables having a two-parameter Weibull distribution with the probability density function of Y_i given by

$$f(y_i;\ \alpha_i,\ \beta_i) = \alpha_i \beta_i y_i^{\beta_i - 1} \exp(-\alpha_i y_i^{\beta_i}),\ \ \alpha_i,\ \beta_i \text{ and } y_i > 0,\ \ i = 1,\ \ldots,\ n. \tag{1}$$

Our interest is to test for no changes in the Weibull parameters against the alternative that there exists an unknown change-point. Accordingly, we formulate the statistical hypotheses as:

$$H_0: \alpha_1 = \cdots = \alpha_n = \alpha \text{ and } \beta_1 = \cdots = \beta_n = \beta, \text{ against } H_a: \exists\tau,\ 1 \leqslant \tau \leqslant n-1;\ \alpha_1 = \cdots = \alpha_\tau \neq \alpha_{\tau+1} = \cdots = \alpha_n \text{ and } \beta_1 = \cdots = \beta_\tau \neq \beta_{\tau+1} = \cdots = \beta_n, \tag{2}$$

where τ is the unknown change-point. When $\tau = t$ (a fixed value) is known, the generalized likelihood ratio statistic for testing H_0 against H_a is given by

$$\Lambda_{t,n} = \frac{\left\{(\tilde{\alpha}\tilde{\beta})^t \prod_{i=1}^{t} y_i^{\tilde{\beta}-1} \exp\left(-\tilde{\alpha}\sum_{i=1}^{t} y_i^{\tilde{\beta}}\right)\right\}\left\{(\alpha^*\beta^*)^{n-t} \prod_{i=t+1}^{n} y_i^{\beta^*-1} \exp\left(-\alpha^*\sum_{i=t+1}^{n} y_i^{\beta^*}\right)\right\}}{\left\{(\hat{\alpha}\hat{\beta})^n \prod_{i=1}^{n} y_i^{\hat{\beta}-1} \exp\left(-\hat{\alpha}\sum_{i=1}^{n} y_i^{\hat{\beta}}\right)\right\}} \tag{3}$$

where $\tilde{\alpha}$, $\tilde{\beta}$ are the mles based on $Y_1, \ldots, Y_t$; α^*, β^* are the mles based on $Y_{t+1}, \ldots, Y_n$ and $\hat{\alpha}$, $\hat{\beta}$ are the mles based on $Y_1, \ldots, Y_n$. The mles $\hat{\alpha}$ and $\hat{\beta}$ may be computed by solving the system of non-linear likelihood equations given by

$$\begin{aligned} \frac{1}{\hat{\alpha}} - \frac{1}{n}\sum_{i=1}^{n} Y_i^{\hat{\beta}} &= 0 \\ \frac{1}{\hat{\beta}} + \frac{1}{n}\sum_{i=1}^{n} \log Y_i - \frac{\hat{\alpha}}{n}\sum_{i=1}^{n} Y_i^{\hat{\beta}} \log Y_i &= 0. \end{aligned} \tag{4}$$

Similar systems of equations may be applied to compute both $\tilde{\alpha}$, $\tilde{\beta}$ and α^*, β^*.

When τ is unknown, the log likelihood ratio statistic may be obtained as

$$Q_n = \max_{1 \leqslant t \leqslant n-1} (2 \log \Lambda_{t,n}). \tag{5}$$

One rejects the null hypothesis H_0 for large values of Q_n. The finite sample null distribution of Q_n is quite complicated. Its asymptotic null distribution, however, is tractable and the result is stated below with its proof appearing in the Appendix:

$$\lim_{n\to\infty} P(\alpha(\log n)Q_n^{1/2} \leqslant x + b(\log n)) = \exp(-2\mathrm{e}^{-x}),\ x \in \mathcal{R}, \tag{6}$$

where $a(s) = (2\log s)^{1/2}$, and $b(s) = 2\log s + \log\log s$. In our experience, the above limiting distribution provides better accuracy for large values of the sample size n.

3. MAXIMUM LIKELIHOOD ESTIMATION OF THE CHANGE-POINT

We now consider the problem of estimating the unknown change-point through the principle of maximum likelihood estimation (mle). The method has been previously considered by Hinkley (1970, 1972), Worsley (1986), Bhattacharya (1987, 1994), and recently by Jandhyala and Fotopoulos (1998, 1999), Fotopoulos and Jandhyala (1998) and Jandhyala *et al.* (1998). Other approaches to the estimation problem may be found in Cobb (1978), Siegmund (1988), and Rukhin (1994).

In his pioneering work, Hinkley (1970, 1972) derived the asymptotic distribution of the mle of a change-point in a single parameter when the parameter is both known and unknown. Importantly, he found that the distributions of the mle were asymptotically identical when the parameter was known or unknown. Recently, Jandhyala and Fotopoulos (1998, 1999) derived a computationally efficient algorithm for approximating the asymptotic probability distribution of the mle. They also derived sharp upper and lower bounds for these asymptotic probabilities. The bounds as well as the two approximations are applicable widely including members of exponential family, and were implemented to both normal and exponential cases. Furthermore, they established exponential rate of convergence for the probability distribution of the mle from finite samples to the case of infinite samples. Fotopoulos and Jandhyala (1998) derived the exact computable form of the asymptotic distribution of the mle for the exponential case, and Jandhyala *et al.* (1998) addressed the problems of detection and estimation of a change in the variability of a process.

Although the Weibull distribution involves more than one parameter, the basic approach derived for the single parameter case by Hinkley (1970, 1972) and Jandhyala and Fotopoulos (1998, 1999) may also be adapted for the multi-parameter case. It may be noted that the asymptotic equivalence result of Hinkley (1972) holds in the multi-parameter case also as long as the point estimators of the parameters both before and after the change-point are consistent. The regularity conditions of Gombay and Horvath (1994a) stated in the Appendix are sufficient for the consistency to hold. We have already demonstrated in the Appendix that the Weibull distribution satisfies these regularity conditions. Thus, our plan is to develop an algorithmic procedure for determining the asymptotic distribution of the mle for the change-point in the Weibull case assuming the parameters to be known both before and after the unknown change-point. The Weibull distribution is not a member of the exponential family and we are not aware of works in the literature that compute the asymptotic distribution for a member outside of the exponential family.

The asymptotic procedure that we develop provides an approximation as well as both upper and lower bounds for the asymptotic distribution of the mle. In our presentation, the development of the algorithm progresses simultaneously along with its application to the data on extreme temperatures in Uppsala, Sweden. The algorithm, however, may be applied to any arbitrary set of changing Weibull parameters. For purposes of estimating the change-point, we consider here the case of changes occurring in both the Weibull parameters. Among the other two situations, the case of a change in β alone is relatively easier and is not considered in the paper. On the other hand, the case where a change occurs in α alone is equivalent to that of estimating the change-point in an exponential parameter. It may be recalled that the exponential case has been fully discussed in Fotopoulos and Jandhyala (1998) and Jandhyala and Fotopoulos (1999).

Let (α_0, β_0) and (α_1, β_1), $\alpha_0 \neq \alpha_1$, $\beta_0 \neq \beta_1$, be the Weibull parameters before and after the unknown change-point τ, $\tau \in \{1, 2, \ldots, n-1\}$. For fixed known values of (α_0, β_0) and (α_1, β_1) the mle of τ may be obtained as

$$\hat{\tau}_n = \underset{1 \leqslant j \leqslant n-1}{\arg\max} \sum_{i=1}^{j} W_i, \tag{7}$$

where $W_i = \log\,(f(Y_i;\,\alpha_0,\,\beta_0)/f(Y_j;\,\alpha_1,\,\beta_1))$ with f representing the Weibull density function as denoted in (1). When (α_0, β_0) and (α_1, β_1) are unknown, first the induced likelihood function $\tilde{L}(\tau)$ is given by

$$\tilde{L}(\tau) = \max_{(\alpha_0,\,\beta_0)} \prod_{i=1}^{\tau} f(y_i;\;\alpha_0,\;\beta_0) \max_{(\alpha_1,\,\beta_1)} \sum_{i=\tau+1}^{n} f(y_i,\;\alpha_1,\;\beta_1). \tag{8}$$

The maximum likelihood estimate $\tilde{\tau}_n$ in this case is:

$$\tilde{\tau}_n = \mathop{\arg\max}_{1 \leqslant j \leqslant n-1} \tilde{L}(j). \tag{9}$$

Since $\hat{\tau}_n$ and $\tilde{\tau}_n$ have the same asymptotic distribution, henceforth, we shall work with the distribution of $\hat{\tau}_n$ only. Our interest is to compute the asymptotic probability distribution of $\hat{\tau}_n$, namely that of $\hat{\tau}_\infty$ obtained by letting both $\tau \to \infty$ and $n - \tau \to \infty$. It is, however, more convenient to work with $\hat{\tau}_n - \tau$ instead of $\hat{\tau}_n$ whereby we have

$$\hat{\tau}_n - \tau = \mathop{\arg\max}_{-\tau+1 \leqslant j \leqslant n-\tau-1} \sum_{i=1}^{\tau+j} W_i = \mathop{\arg\max}_{-\tau+1 \leqslant j \leqslant n-\tau-1} \xi(j), \tag{10}$$

where

$$\xi(j) = \sum_{i=1}^{\tau+j} W_i - \sum_{i=1}^{\tau} W_i = \begin{cases} \sum_{i=1}^{j} X_i^* = S_j^*, & j > 0 \\ 0, & j = 0 \\ \sum_{i=1}^{-j} X_i = S_{-j}, & j < 0, \end{cases}$$

with $X_i = -W_{-\nu-i+1}$, for $1 \leqslant i \leqslant \tau - 1$ and $X_j^* = W_{\tau+1}$, for $1 \leqslant i \leqslant n - \tau$. The random walks $S = \{S_n{:}n \geqslant 0\}$ and $S^* = \{S_n^* : n \geqslant 0\}$ are independent but not necessarily identical. Furthermore, both walks have negative means and will eventually drift to $-\infty$. The algorithm that we develop for computing the asymptotic distribution of $\hat{\tau}_n - \tau$ depends extensively on the two random walks S and S^*. Let X and X^* represent the initial random variables associated with S and S^* respectively.

In the sequel, we shall require the probability distributions as well as the moment generating functions of both X and X^*. First, we express X and X^* in terms of U, a uniform [0,1] random variable.

$$X = \frac{\beta_1 - \beta_0}{\beta_0} \log\left(\alpha_0^{-1} \log\frac{1}{1-U}\right) + \log\frac{1}{1-U} - \alpha_1\left(\alpha_0^{-1}\log\frac{1}{1-U}\right)^{\beta_1/\beta_0} + \log\frac{\alpha_1\beta_1}{\alpha_0\beta_0} \tag{11}$$

$$X^* = \frac{\beta_0 - \beta_1}{\beta_1} \log\left(\alpha_1^{-1} \log\frac{1}{1-U}\right) + \log\frac{1}{1-U} - \alpha_0\left(\alpha_1^{-1}\log\frac{1}{1-U}\right)^{\beta_0/\beta_1} + \log\frac{\alpha_0\beta_0}{\alpha_1\beta_1}. \tag{12}$$

It may be noted that X and X^* are dual to each other in the sense that one may be obtained from the other by interchanging the subscripts in the parameters (α_0, β_0) and (α_1, β_1). Some tedious calculations will show that X and X^* have the following moment generating functions.

$$m(s) = \left(\frac{\beta_1}{\beta_0}\right)^{s-1}\left(\frac{\alpha_1^{\beta_0/\beta_1}}{\alpha_0}\right)^{s-1} s^{-s+\beta_0/\beta_1(s-1)} \sum_{j=0}^{\infty} \frac{\left(-\frac{1-s}{s^{\beta_0/\beta_1}}\right)^j}{j!} \times \left(\frac{\alpha_1^{\beta_0/\beta_1}}{\alpha_0}\right)^{-j} \Gamma\left(\frac{\frac{\beta_1-\beta_0}{\beta_0}s+j+1}{\beta_1/\beta_0}\right) \tag{13}$$

$$m^*(s) = \left(\frac{\beta_0}{\beta_1}\right)^{s-1}\left(\frac{\alpha_0^{\beta_1/\beta_0}}{\alpha_1}\right)^{s-1} s^{-s+\beta_1/\beta_0(s-1)} \sum_{j=0}^{\infty} \frac{\left(-\frac{1-s}{s^{\beta_1/\beta_0}}\right)^j}{j!} \times \left(\frac{\alpha_1^{\beta_1/\beta_0}}{\alpha_0}\right)^{-j} \Gamma\left(\frac{\frac{\beta_0-\beta_1}{\beta_1}s+j+1}{\beta_0/\beta_1}\right). \tag{14}$$

We shall now propose methods for computing upper and lower bounds as well as two approximations for the asymptotic distribution of $\hat{\tau}_n - \tau$. The methodology follows somewhat analogously to the treatment contained in Jandhyala and Fotopoulos (1999). Accordingly, we first require the following quantities all related to X and X^*. However, due to the duality between X and X^*, we omit the expressions for the '*' counterpart where no confusion arises.

$$\vartheta = \sup\{s \in \mathcal{R} : m(s) \leqslant 1\},$$

$$h(t) = \frac{P(X > t)}{\int_t^\infty \exp(\vartheta(x-t))\mathrm{d}P(X \leqslant x)}, \quad a_1 = \sup_{t>0} h(t) \text{ and } a_2 = \inf_{t>0} h(t)$$

$$b_n = P(S_n > 0), \; d_n = E[S_n I(S_n > 0)], \; \tilde{b}_n(\vartheta^*) = E[\mathrm{e}^{-\vartheta^* S_n} I(S_n > 0)], \; n \geqslant 1$$

$$B(s) = \sum_{n=1}^{\infty} s^n b_n/n, \; 0 \leqslant s \leqslant 1; \; p = 1 - \mathrm{e}^{-B(1)}; \; \mu = \sum_{n=1}^{\infty} \mathrm{d}_n/n$$

and the sequences $\{q_n; n \geqslant 0\}$, $\{\tilde{u}_n(\vartheta^*); n \geqslant 0\}$ obtained from the iterative procedures:

$$q_0 = 1 \text{ and } nq_n = \sum_{j=0}^{n-1} b_{n-j} q_j, \; n \geqslant 1; \; \tilde{u}_0(\vartheta^*) = 1 \text{ and } n\tilde{u}_n(\vartheta^*) = \sum_{j=0}^{n-1} \tilde{b}_{n-j}(\vartheta^*)\tilde{u}_j(\vartheta^*), \; n \geqslant 1. \tag{15}$$

Further, denote $A_j(\omega) = e^{-B^*(1)}\{q_j^* - \omega\tilde{u}_j^*(\vartheta)\}$ and $A_j^*(\omega) = e^{-B(1)}\{q_j - \omega\tilde{u}_j(\vartheta^*)\}$, $j \in Z^+$, where ω may assume one of several parameters. As demonstrated in Jandhyala and Fotopoulos (1999), the asymptotic probabilities then admit the bounds stated in I below, and the two approximations stated in II and III.

(I) $A_j(\alpha_1) \leqslant P(\hat{\tau}_\infty - \tau = j) \leqslant A_j(\alpha_2)$, $A_j^*(\alpha_1^*) \leqslant P(\hat{\tau}_\infty - \tau = -j) \leqslant A_j^*(\alpha_2^*)$, $j \in Z^+$

(II) $A_j(p) \cong P(\hat{\tau}_\infty - \tau = j)$, $A_j^*(p^*) \cong P(\hat{\tau}_\infty - \tau = -j)$, $j \in Z^+$
(III) $A_j(\mu\vartheta) \cong P(\hat{\tau}_\infty - \tau = j)$, $A_j^*(\mu^*\vartheta^*) \cong P(\hat{\tau}_\infty - \tau = -j)$, $j \in Z^+$.

In some applications, it may happen that $\vartheta = \vartheta^* = 1$, in which case the approximations II and III would be identical.

Computation of the bounds in I and the approximations in II and III primarily requires that one is able to compute ϑ, $h(t)$, $\{b_n\}$, $\{d_n\}$ and $\{\tilde{b}_n(\vartheta^*)\}$ and their '*' counterparts. The remaining parameters and sequences may then be obtained routinely as functions of the above. We shall now state some remarks regarding the computation of ϑ, $h(t)$, $\{b_n\}$, $\{d_n\}$ and $\{\tilde{b}_n(\vartheta^*)\}$.

Computation of ϑ. The moment generating function $m(s)$ given in (13) provides the value for ϑ. It may, however, be noted that (13) is quite complicated in s, and at times may be sensitive for computations. In such cases, we suggest the evaluation of $m(s)$ based on simulations. The simulations are quite straightforward given that X as in (11) is expressed conveniently in terms of U, the uniform [0,1] random variable.
Computation of $h(t)$. This requires the computation of $P(X > t)$ and

$$\int_t^\infty \exp(\vartheta(x - t))\mathrm{d}P(X \leqslant x).$$

Both of these quantities are complicated for direct computation. Consequently, we found it convenient to obtain $h(t)$ based on simulations. The simulations involve only X and one can obtain any desired level of accuracy by performing sufficiently large number of simulations.
Computation of $\{b_n\}$, $\{d_n\}$ *and* $\{\tilde{b}_n(\vartheta^*)\}$. Here again, direct computations appear intractable and we suggest using simulations when n is small. Otherwise, excellent approximations are available in the literature (Veraverbeke and Teugels 1975). The approximations are easy to apply and in many situations readily provide values of good accuracy when $n = 10$ or even smaller depending on the parameter values. The relevant approximations are stated below:

(i) $b_n \cong c_1 n^{-1/2}\gamma^n$, $c_1 = 1/\eta\nu\sqrt{2\pi}$
(ii) $d_n \cong c_2 n^{-1/2}\gamma^n$, $c_2 = 1/\eta\nu^2\gamma^{1/2}\sqrt{2\pi}$
(iii) $\tilde{b}_n(\vartheta^*) \cong c_3 n^{-1/2}\gamma^n$, $c_3 = 1/\eta(\nu + \vartheta^*)\sqrt{2}\pi$.

where, $\eta = \sqrt{m''(\nu)/\gamma}$, $\gamma = \inf_s m(s)$ and ν is such that $m(\nu) = \inf_s m(s)$.

4. CHANGE-POINT ANALYSIS OF DATA ON EXTREME TEMPERATURES

The process of measuring daily temperature measurements started in Uppsala, Sweden, as early as 1712–1713. Anders Celsius, then Professor of astronomy at the University of Uppsala, showed great interest in collecting temperature measurements and there is almost complete daily temperature data from the year 1739. The measurements for the early period of 1739–1838 were made only three or four times a day and thus were based on discrete recordings. With the installation of a max–min thermometer in 1839, recording of the true minimum was possible through continuous temperature readings. Thus, lower extreme minimum values were to be expected subsequent to the year 1839. This effect may be expected to be particularly large for the minimum temperatures during the summer months, since the lowest temperatures occur very early in the morning at that latitude. The discrete recordings prior to 1839 were, however, made when the sun had been up for several hours. While discussing applications of extreme value

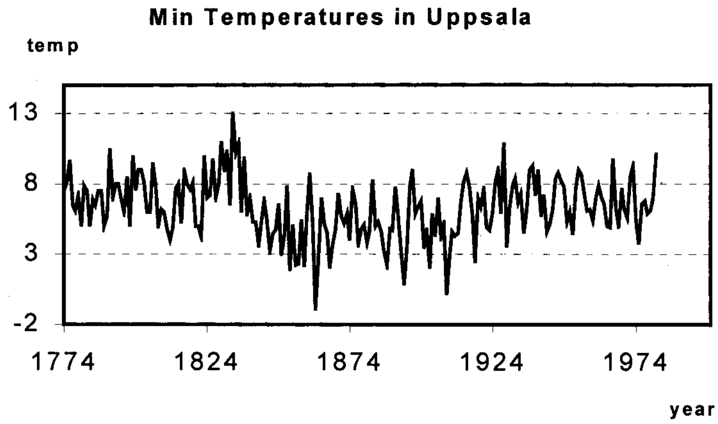

Figure 1. Minimum temperatures in Uppsala, Sweden, 1774–1981

theory, Leadbetter *et al.* (1983, p. 279) reported the minimum temperature data for the month of July beginning from the year 1739. They extracted this data from the original manuscripts compiled by Sverker Hellstrom. The early years of data, however, had some missing observations and complete data was reported beginning from the year 1774. Figure 1 thus represents the monthly minimum temperatures in Uppsala, Sweden for the month of July between the years 1774 and 1981, consisting of 208 observations. Due to the inaccessibility of the original numerical values, we also present the data in Table I as read from Figure 15.1.1(b) of Leadbetter *et al.* (1983, p. 279).

Table I. Minimum temperatures in Uppsala, 1774–1981 (numbers are read from left to right)

9·5	10·2	11·7	8·5	8·1	9·2	7·0	9·9	9·5	7·0
8·9	8·5	9·5	9·5	7·0	7·7	12·5	9·0	10·0	10·0
8·9	8·0	10·5	7·0	12·0	9·6	11·0	11·0	10·1	8·0
8·0	11·5	9·8	7·0	8·2	8·0	6·8	6·0	6·9	9·7
10·0	7·2	10·9	10·0	9·6	10·1	7·0	7·0	6·2	12·0
9·0	9·2	11·8	9·0	10·0	13·0	11·0	12·4	8·5	15·1
12·2	12·9	8·0	11·9	7·8	9·0	7·3	7·3	5·6	7·3
8·9	7·4	5·2	6·5	6·7	8·6	4·9	6·3	9·9	3·8
7·1	4·2	4·3	7·5	4·1	8·0	10·8	7·1	1·0	4·8
9·0	7·2	6·5	4·0	5·7	6·7	9·3	7·7	7·2	7·9
6·0	9·7	8·6	5·7	6·8	7·1	5·8	6·6	10·3	7·0
7·3	6·6	5·2	4·2	6·8	6·7	9·8	7·7	5·2	2·8
5·4	9·8	11·0	7·9	8·4	8·8	5·5	6·7	4·0	7·9
6·3	9·0	6·2	7·2	2·1	4·9	6·6	6·3	6·5	8·7
10·2	10·8	9·7	7·7	4·4	9·0	8·4	9·7	6·9	6·7
8·0	10·0	11·0	7·9	12·9	5·5	8·3	9·9	10·4	8·7
9·3	6·5	8·3	11·0	11·3	9·2	11·0	7·7	9·2	6·6
7·1	8·2	10·4	10·8	10·2	9·8	7·3	8·0	6·4	9·7
11·0	10·7	9·4	8·1	8·2	7·4	9·0	9·9	9·0	8·6
7·0	6·9	11·8	8·2	7·0	9·7	8·2	7·6	10·5	11·3
7·4	5·7	8·6	8·8	7·9	8·1	9·0	12·1		

The data being monthly minimum temperatures (extreme values), one may expect the Weibull distribution to fit the data well. In this analysis, we are concerned with the dynamic stability of the scale and shape parameters associated with the Weibull distribution. We formulate this dynamic stability as a change-point inferential problem. The formulation begins with the assumption that the monthly minimum temperatures are independent. Subsequently, we will substantiate the validity of the independence assumption as well as that of a Weibull fit to the data. To ensure positivity, we add an arbitrary value (2°C) to all observations.

4.1. Detection of change-points

First, we apply the detection methodology discussed in Section 2 in order to detect unknown changes in the monthly minimum temperature data. Figure 2 is a plot of the log likelihood ratio statistic, $2\log\Lambda_{t,n}$ for $t = 1, \ldots, 208$. From Section 2, it follows that this statistic tests for unknown changes in both α and β, the scale and shape parameters of the underlying Weibull distribution.

This plot shows the maximum at the year 1837 with the statistic being $Q_{208} = 27{\cdot}51$. The sample size $n = 208$ being large, the asymptotic distribution of Q_n given in (6) is applicable. The corresponding p-value $= 0{\cdot}00645$ and thus the analysis clearly identifies a significant change at the year 1837 ($\hat{\tau}_{208} = 64$).

Figure 2 also shows several local maxima around 1900–1930. The occurrence of these local maxima might illustrate the instability of the likelihood ratio or the presence of another change-point between the years 1900 and 1930. This requires further analysis on data for the years 1838–1981. The corresponding statistic value is $Q_{144} = 39{\cdot}62$, with p-value $= 0{\cdot}001$, thus identifying another significant change during the year 1912 ($\hat{\tau}_{144} = 75$). We couldn't fail to notice a diminishing effect of the second change-point at 1912 on the overall statistic Q_{208} with all the 208 observations. For this reason, we considered data between the years 1774 and 1912 where we may anticipate the occurrence of only a single change-point. The corresponding statistic is $Q_{139} = 51{\cdot}75$ with maximum at 1839 ($\hat{\tau}_{139} = 66$) and the p-value $= 0{\cdot}0002$. As expected, this p-value is now much higher than the previous p-value with all the 208 observations. Note that the change-point estimated shifted by two years, from 1937 to the year 1939. Since the data between

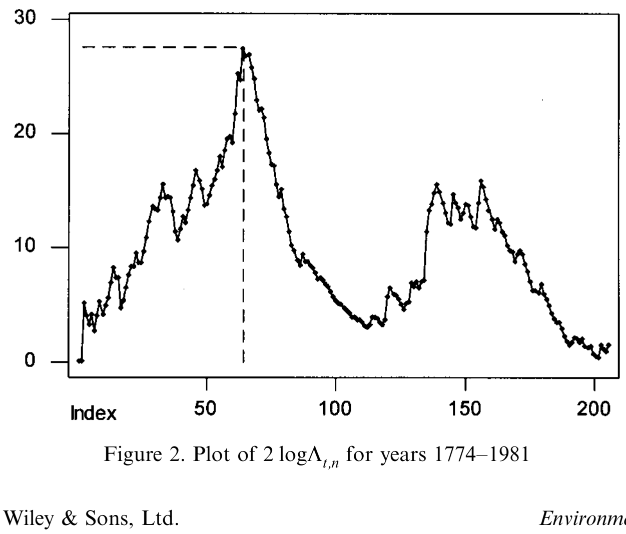

Figure 2. Plot of $2\log\Lambda_{t,n}$ for years 1774–1981

the years 1774 and 1912 is free of the effect of the second change-point, we consider the year 1939 to be the mle for the first change-point. No other significant changes were found in the data.

The first change detected at the year 1839 may be justified by the transition in the method of data collection. We may recall that prior to the year 1839, data was collected discretely based on three or four daily measurements, sometimes at irregular intervals. However, from 1839 onwards, temperatures were measured with the max–min thermometer and thus gave precise values for the extreme measurements. The justification for the second change-point detected at the year 1912 required a careful enquiry. By the end of the 19th century, apparently, the city of Uppsala was transformed from a small educational and religious center to an industrialized transportation hub. Some urbanization effect should therefore be expected around the beginning of the 20th century, thus resulting in higher temperatures from that period onwards.

Before we apply the estimation analysis of Section 3, we present the p-values for Weibull χ^2 goodness-of-fit tests for the three data segments 1774–1839, 1840–1912 and 1913–1981. The respective p-values are 0·1764, 0·5911 and 0·9567. These values indicate the appropriateness of the Weibull model for all the three segments. The evidence is particularly overwhelming for the last two data segments. We also present in Figure 3 the auto-correlation and partial auto-correlation plots for the residuals of the three data segments. The plots clearly show no evidence of serial correlations in the data and thus support the assumption of independence among the time-ordered observations.

4.2. Estimation of the unknown change-points

Here, we implement the estimation methodology presented in Section 3 and calculate the asymptotic probability distributions of the mles of the two unknown change-points. In particular, this will enable us to provide confidence intervals of any desired level for the change-points. Since

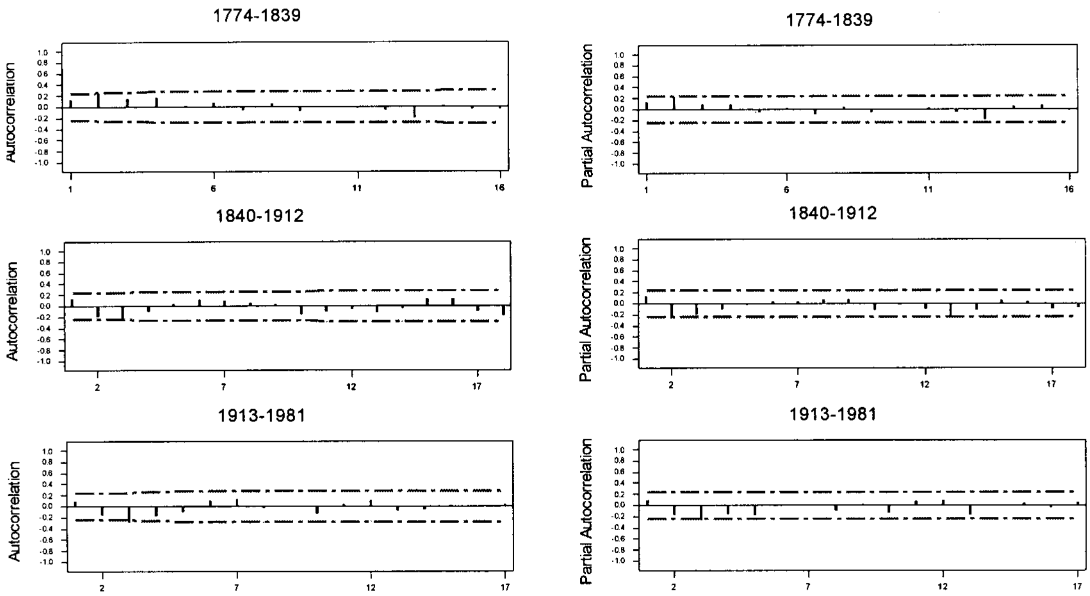

Figure 3. The auto-correlation and partial auto-correlation plots of the three segments

the mle is known to be inconsistent, the information provided by the confidence intervals assumes greater significance.

We fully illustrate the estimation methodology for the data set 1774–1912 with the mle given by $\hat{\tau} = 66$, i.e., the year 1839. First, we present the conditional mles of the Weibull scale and shape parameters corresponding to before and after the mle $\hat{\tau} = 66$. Solving the two likelihood equations in (4), we get $\hat{\alpha}_{0|\hat{\tau}=66} = 0{\cdot}0000057$, $\hat{\beta}_{0|\hat{\tau}=66} = 5{\cdot}19613$ and $\hat{\alpha}_{1|\hat{\tau}=66} = 0{\cdot}000429$, $\hat{\beta}_{1|\hat{\tau}=66} = 3{\cdot}86486$. Basing upon the asymptotic distributional equivalence result of Hinkley (1972), we assume the above conditional estimates to provide known values for the respective parameters. Thus, we let $\alpha_0 = 0{\cdot}0000057$, $\beta_0 = 5{\cdot}19613$ and $\alpha_1 = 0{\cdot}000429$, $\beta_1 = 3{\cdot}86486$. These parameter values enable us to generate the random variables X and X^* given by equations (11) and (12).

The estimation methodology discussed in Section 3 requires that we first compute ϑ, α_1, α_2, $\{b_n\}$, $\{\tilde{b}_n(\vartheta^*)\}$, $\{d_n\}$, $B(1)$, p and μ. We also need to compute the '*' counterparts of the above quantities. Furthermore, in order to implement the large scale approximations suggested by Veraverbeke and Teugels (1975), we need to compute γ, ν and η and also their '*' counterparts. These computations begin by first computing the mgfs $m(s)$ and $m^*(s)$ expressed respectively in (13) and (14). Since these expressions are complicated, we found it convenient to evaluate the two mgfs based on simulations. The two mgfs computed are presented in Figure 4.

The computations result in $\vartheta = 1{\cdot}00083$, $\gamma = 0{\cdot}798595$, $\nu = 0{\cdot}46303$, $\eta = 1{\cdot}34487$ and $\vartheta^* = 1{\cdot}00588$, $\gamma^* = 0{\cdot}798358$, $\nu^* = 0{\cdot}53852$, $\eta^* = 1{\cdot}34071$. Next we evaluate the functions $h(t)$ and $h^*(t)$ (see Figure 5) to get the values for α_1, α_2 and α_1^*, α_2^*.

Accordingly, we have $\alpha_1 = 1$, $\alpha_2 = 0$ and $\alpha_1^* = 1$, $\alpha_2^* = 0$. The quantities $\{b_n\}$, $\{\tilde{b}_n(\vartheta^*)\}$ and their '*' counterparts may now be computed. The first 10 values in each of these sequences have been computed based on 200,000 simulations. Subsequent values were all computed using the

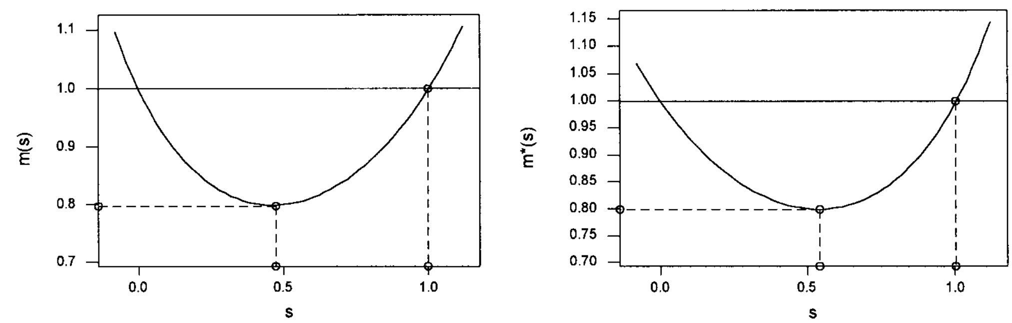

Figure 4. The moment generating functions $m(s)$ and $m^*(s)$

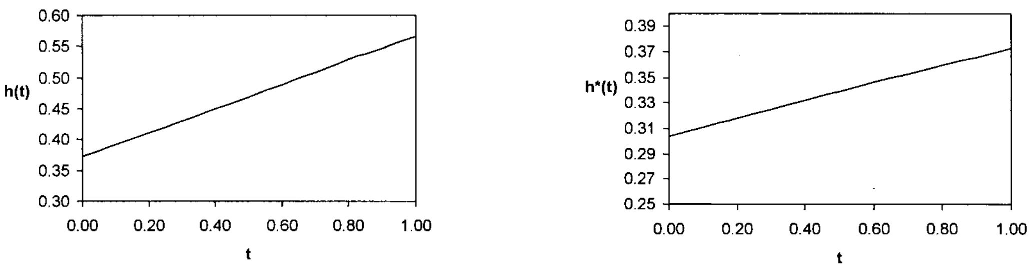

Figure 5. The functions $h(t)$ and $h^*(t)$

Table II. Computed values for various sequences

n	b_n	b_n^*	$\tilde{b}_n(\vartheta^*)$	$\tilde{b}_n^*(\vartheta)$	q_n	q_n^*	$\tilde{u}_n(\vartheta^*)$	$\tilde{u}_n^*(\vartheta)$
1	0·2936	0·2135	0·1520	0·1155	0·2936	0·2135	0·1520	0·1155
2	0·1942	0·1515	0·0876	0·0737	0·1402	0·0985	0·0553	0·0435
3	0·1390	0·1087	0·0583	0·0511	0·0791	0·0540	0·0267	0·0215
4	0·1001	0·0788	0·0402	0·0349	0·0478	0·0321	0·0145	0·0116
5	0·0738	0·0608	0·0283	0·0262	0·0304	0·0207	0·0084	0·0071
6	0·0556	0·0464	0·0217	0·0197	0·0201	0·0137	0·0054	0·0045
7	0·0414	0·0345	0·0152	0·0140	0·0135	0·0092	0·0034	0·0028
8	0·0324	0·0266	0·0114	0·0110	0·0094	0·0063	0·0022	0·0019
9	0·0249	0·0206	0·0088	0·0080	0·0066	0·0044	0·0015	0·0013
10	0·0196	0·0166	0·0068	0·0066	0·0047	0·0032	0·0011	0·0009
11	0·0163	0·0140	0·0051	0·0049	0·0035	0·0024	0·0007	0·0006
12	0·0124	0·0107	0·0039	0·0037	0·0026	0·0018	0·0005	0·0005
13	0·0095	0·0082	0·0030	0·0029	0·0019	0·0013	0·0004	0·0003
14	0·0073	0·0063	0·0023	0·0022	0·0014	0·0009	0·0003	0·0002
15	0·0057	0·0049	0·0018	0·0017	0·0010	0·0007	0·0002	0·0002
16	0·0044	0·0038	0·0014	0·0013	0·0007	0·0005	0·0001	0·0001
17	0·0034	0·0029	0·0011	0·0010	0·0006	0·0004	0·0001	0·0001
18	0·0026	0·0023	0·0008	0·0008	0·0004	0·0003	0·0001	0·0001
19	0·0020	0·0018	0·0006	0·0006	0·0003	0·0002	0·0001	0·0001
20	0·0016	0·0014	0·0005	0·0005	0·0002	0·0002	0·0000	0·0000
21	0·0012	0·0011	0·0004	0·0004	0·0002	0·0001	0·0000	0·0000
22	0·0010	0·0008	0·0003	0·0003	0·0001	0·0001	0·0000	0·0000
23	0·0008	0·0006	0·0002	0·0002	0·0001	0·0001	0·0000	0·0000
24	0·0006	0·0005	0·0002	0·0002	0·0001	0·0001	0·0000	0·0000
25	0·0005	0·0004	0·0001	0·0001	0·0001	0·0000	0·0000	0·0000

large sample approximations suggested in Section 3. In the case of this example, we found the large sample approximations to be quite accurate from as small as $n = 5$ onwards for all the sequences. The sequences $\{q_n\}$, $\{\tilde{u}_n(\vartheta^*)\}$ and their '*' counterparts may now be computed based on the recursive relations expressed in (15). Table II provides the values for all of the sequences up to $n = 25$.

Recall that we have $\vartheta = 1{\cdot}00083$ and $\vartheta^* = 1{\cdot}00588$. Since these values are close to unity, it follows from our discussion in Section 3 that the approximations II and III for the asymptotic distribution of the mle will be identical. Thus, in Table III, we now provide the lower bound, approximation II and the upper bound for the asymptotic probabilities, $P(\hat{\tau}_\infty - \tau = j), j = -20, -19, \ldots, 20$. In the table, the 'Sum' refers to sum of all the significant terms (up to the seventh decimal place) in each column. Accordingly, the 'Sum' is based on terms for $j = -30, -29, \ldots, 30$ even though the probabilities presented range between $j = -20, -19, \ldots, 20$. We may note that the bounds are quite tight and that the approximation works out to be extremely accurate. The shortest 95% confidence interval for the unknown change-point, based on the probabilities from the approximation column, is found to be 1839 ± 5, i.e., the years 1834, …, 1844.

The asymptotic probability distribution of the second change-point detected at 1912 has also been computed implementing the above estimation procedure. Table IV provides the corresponding asymptotic probability distribution. The 95% confidence interval for the mle in this case is found to be 1912 ± 7, i.e., the years 1905, …, 1919. For this data segment, the scale and shape parameters after the change-point were found to be 0·0000018513 and 5·85506, respectively.

Table III. Probability distribution of $\hat{\tau}_{\infty} - \tau$

j	Lower bound	Approximation	Upper bound
−20	0·00011	0·00013	0·00014
−19	0·00015	0·00017	0·00018
−18	0·00020	0·00023	0·00025
−17	0·00027	0·00031	0·00033
−16	0·00036	0·00042	0·00045
−15	0·00049	0·00057	0·00060
−14	0·00066	0·00077	0·00082
−13	0·00090	0·00105	0·00112
−12	0·00123	0·00144	0·00154
−11	0·00167	0·00197	0·00211
−10	0·00222	0·00265	0·00285
−9	0·00306	0·00368	0·00397
−8	0·00433	0·00524	0·00566
−7	0·00611	0·00749	0·00814
−6	0·00886	0·01108	0·01211
−5	0·01326	**0·01673**	0·01834
−4	0·02012	**0·02608**	0·02885
−3	0·03158	**0·04256**	0·04766
−2	0·05116	**0·07394**	0·08453
−1	0·08537	**0·14792**	0·17699
0	0·41161	**0·41161**	0·41161
1	0·06691	**0·11446**	0·14577
2	0·03757	**0·05547**	0·06726
3	0·02217	**0·03104**	0·03687
4	0·01399	**0·01878**	0·02192
5	0·00928	**0·01220**	0·01412
6	0·00629	0·00815	0·00937
7	0·00433	0·00549	0·00626
8	0·00300	0·00380	0·00432
9	0·00216	0·00269	0·00303
10	0·00156	0·00194	0·00220
11	0·00122	0·00148	0·00166
12	0·00090	0·00108	0·00121
13	0·00066	0·00079	0·00088
14	0·00048	0·00058	0·00064
15	0·00036	0·00042	0·00047
16	0·00026	0·00031	0·00035
17	0·00020	0·00023	0·00026
18	0·00015	0·00017	0·00019
19	0·00011	0·00013	0·00014
20	0·00008	0·00010	0·00011
Sum	0·815359	1·01533	1·125264

The confidence interval estimates in both cases provide much needed information. It may be noted that the confidence interval for the second change-point includes the early years of the 20th century. History suggests that rapid industrial development took place in Uppsala during that period. This rapid industrialization should have had subsequent effect on the temperatures in and around Uppsala, Sweden.

Table IV. Probability distribution of $\hat{\tau}_{\infty} - \tau$ (second change-point)

j	Lower bound	Approximation	Upper bound
−15	0·00128	0·00148	0·00163
−14	0·00161	0·00188	0·00206
−13	0·00204	0·00238	0·00262
−12	0·00257	0·00302	0·00334
−11	0·00321	0·00380	0·00423
−10	0·00374	0·00452	0·00509
−9	0·00488	0·00597	0·00675
−8	0·00636	0·00784	0·00889
−7	0·00841	**0·01045**	0·01191
−6	0·01141	**0·01438**	0·01652
−5	0·01574	**0·02010**	0·02324
−4	0·02228	**0·02914**	0·03408
−3	0·03195	**0·04347**	0·05175
−2	0·04768	**0·06953**	0·08524
−1	0·07398	**0·12610**	0·16359
0	0·33306	**0·33306**	0·33306
1	0·07226	**0·12398**	0·16259
2	0·04587	**0·06716**	0·08305
3	0·03058	**0·04180**	0·05017
4	0·02115	**0·02782**	0·03279
5	0·01510	**0·01941**	0·02263
6	0·01102	**0·01391**	0·01607
7	0·00812	**0·01010**	0·01158
8	0·00611	0·00756	0·00864
9	0·00462	0·00566	0·00643
10	0·00353	0·00427	0·00482
11	0·00305	0·00363	0·00406
12	0·00244	0·00288	0·00321
13	0·00194	0·00227	0·00252
14	0·00153	0·00179	0·00198
15	0·00122	0·00142	0·00156
Sums	0·80840	1·02185	1·17824

APPENDIX

Proof of equation (6): Gombay and Horvath (1994a) established the asymptotic null distribution of the log likelihood ratio statistic under regularity conditions C1–C7. We refer the reader to their paper for statements of conditions C1–C7. Below, we briefly demonstrate the applicability of these regularity conditions for the two parameter Weibull distribution.

Let $g(y; \alpha, \beta) = \log f(y; \alpha, \beta)$ and let

$$g_{\alpha^m\beta^k} = \frac{\partial^{m+k}}{\partial\alpha^m \partial\beta^k} g, \quad m, k = 0, 1, \ldots$$

denote the respective partial derivatives of g with respect to α and β.

C1: The uniqueness of the maximum likelihood estimators is shown in Farnum and Booth (1997).

C2: The existence and continuity of $g_{\alpha^m\beta^k}$, $1 \leqslant m + k \leqslant 3$, may be verified easily.

C3: The boundedness and integrability of $g_{\alpha^m\beta^k}$, $1 \leqslant k \leqslant 3$ also follow in a straightforward manner.

C4: First, one shows that $g_\alpha(y;\ \alpha,\ \beta) = 1/\alpha - y^\beta$ and $g_\beta(y;\ \alpha,\ \beta) = 1/\beta + \log y - \alpha y^\beta \log y$. Then, it can be shown that $E[Y^\beta] = 1/\alpha$ and $E[\log Y] = -1/\beta + \alpha E[Y^\beta \log y]$. It follows then that $E[g_\alpha(y; \alpha, \beta)] = E[g_\beta(y; \alpha, \beta)] = 0$.

C5: We need to show: (i) $E[g_\alpha^2] = -E[g_{\alpha^2}]$; (ii) $E[g_\alpha g_\beta] = -E[g_{\alpha\beta}]$; (iii) $E[g_\beta^2] = -E[g_{\beta^2}]$.

As for the next part, that the inverse of the information matrix exists and its elements are continuous follows from the forms of the above expectations. We only provide details for showing (iii). The details for showing (i) and (ii) are relatively easier. First note that

$$g_\beta^2 = \frac{1}{\beta^2} + \frac{2}{\beta}\log y + \log^2 y - \frac{2\alpha}{\beta} y^\beta \log y - 2\alpha y^\beta \log^2 y + \alpha^2 y^{2\beta} \log^2 y, \text{ and}$$
$$g_{\beta^2} = -\left(\frac{1}{\beta^2} + \alpha y^\beta \log^2 y\right).$$

The following non-trivial expressions then prove (iii):

$$E[Y^\beta \log Y] = \frac{1}{\alpha\beta} + \frac{1}{\alpha}E[\log Y];\ \ E[Y^{2\beta} \log Y] = \frac{2}{\alpha^2\beta} + \frac{1}{\alpha^2}E[\log Y] + \frac{1}{\alpha}E[Y^\beta \log Y];$$
$$E[Y^\beta \log^2 Y] = \frac{2}{\alpha\beta}E[\log Y] + \frac{1}{\alpha}E[\log^2 Y];$$
$$E[Y^{2\beta} \log^2 Y] = \frac{2}{\alpha^2\beta}E[\log Y] + \frac{1}{\alpha^2}E[\log^2 Y] + \frac{2}{\alpha\beta}E[Y^\beta \log Y] + \frac{1}{\alpha}E[Y^\beta \log^2 Y].$$

Conditions C6 and C7 follow upon showing that $E[Y^\rho \log^m Y] < \infty$, $\rho \in R^+$ and $m \in \mathcal{N}$. The finiteness of this expectation, however, follows from

$$\int_0^\infty x^{\rho-1} e^{-\sigma x} \log^m x \,\mathrm{d}x = \frac{\partial^m}{\partial v^m}(\sigma^{-\rho}\Gamma(\rho)) < \infty,$$

where $\rho, \sigma > 0$ and $m \in \mathcal{N}$. The above equation may be found in Gradshteyn and Ryzhik (1980, p. 578).

REFERENCES

Bhattacharya, P. K. (1987). 'Maximum likelihood estimation of a change-point in the distribution of independent random variables: General multi-parameter case'. *Journal of Multivariate Analysis* **23**, 183–208.

Bhattacharya, P. K. (1994). 'Some aspects of change-point analysis'. In *Change-point Problems*, eds. E. Carlstein, H. G. Muller and D. Siegmund. 000: IMS Monograph, pp. 28–56.

Brillinger, D. R. (1994). 'Some river wavelets'. *Environmetrics* **5**, 211–220.

Brillinger, D. R. (1997). 'Random process methods and environmental data: The 1996 Hunter Lecture'. *Environmetrics* **8**, 269–281.

Cobb, G. W. (1978). 'The problem of the Nile: Conditional solution to a change-point problem'. *Biometrika* **62**, 243–251.

Davis, R. A., Huang, D. and Yao, Y. C. (1995). 'Testing for a change in the parameter values and order of an auto-regressive model'. *Annals of Statistics* **23**, 282–304.

Essenwanger, O. M. (1976). *Applied Statistics in Atmospheric Science*. Amsterdam: Elsevier.

Farnum, N. and Booth, P. (1997). 'Uniqueness of maximum-likelihood estimators of the 2-parameter Weibull distribution'. *IEEE Transactions on Reliability* **46**, 523–525.

Fotopoulos, S. B. and Jandhyala, V. K. (1998). 'Maximum likelihood estimation of a change-point for exponentially distributed random variables' (submitted).

Glantz, M. H. (1987). *Drought and Hunger in Africa: Denying Famine a Future*. Cambridge: Cambridge University Press.

Gombay, E. and Horvath, L. (1990). 'Asymptotic distributions of maximum likelihood tests for changing in the mean'. *Biometrika* **77**, 411–415.

Gombay, E. and Horvath, L. (1994a). 'Limit theorems for change in linear regression'. *Journal of Multivariate Analysis* **48**, 43–69.

Gombay, E. and Horvath, L. (1994b). 'An application of the maximum likelihood test to the change-point problem'. *Stochastic Processes and Their Applications* **50**, 161–171.

Gombay, E. and Horvath, L. (1996a). 'On the rate of approximation for maximum likelihood tests in change-point models'. *Journal of Multivariate Analysis* **56**, 120–152.

Gombay, E. and Horvath, L. (1996b). 'Approximations for the time of change and the power function in change-point models'. *Journal of Statistical Planning and Inference* **52**, 43–66.

Gombay, E. and Horvath, L. (1997). 'An application of the likelihood method to change-point detection'. *Environmetrics* **8**, 459–467.

Gradshteyn, I. S. and Ryzhik, I. M. (1980). *Tables of Integrals, Series and Products*. New York: Academic Press.

Haccou, P., Meeks, E. and Geer, S. (1988). 'The likelihood ratio test for the change-point problem for exponentially distributed random variables'. *Stochastic Processes and Their Applications* **27**, 121–139.

Hawkins, D. M. (1977). 'Testing a sequence of observations for a shift in location'. *Journal of American Statistical Association* **72**, 180–186.

Hinkley, D. V. (1970). 'Inference about the change-point in a sequence of random variables'. *Biometrika* **57**, 1–17.

Hinkley, D. V. (1972). 'Time order classification'. *Biometrika* **59**, 509–523.

Horvath, L. (1993a). 'The maximum likelihood method for testing changes in the parameters of normal observations'. *Annals of Statistics* **21**, 671–680.

Horvath, L. (1993b). 'Change in Auto-regressive processes'. *Stochastic Processes and Their Applications* **44**, 221–242.

James, B., James, K. L. and Siegmund, D. (1987). 'Tests for a change-point'. *Biometrika* **74**, 71–83.

Jandhyala, V. K. and Fotopoulos, S. B. (1998). 'The rate of convergence of the maximum likelihood estimator of a change-point' (submitted).

Jandhyala, V. K. and Fotopoulos, S. B. (1999). 'Capturing the distributional behavior of the maximum likelihood estimate of a change-point'. *Biometrika* **86**, 129–140.

Jandhyala, V. K., Fotopoulos, S. B. and Hawkins, D. M. (1998). 'Detection and estimation of abrupt changes in the variability of a process' (submitted).

Jaruskova, D. (1994). 'Applications of change-point detection to ecology'. Preprint.

Jaruskova, D. (1997). 'Some problems with application of change-point detection methods to environmental data'. *Environmetrics* **8**, 469–483.

Katz, R. W. and Brown, B. G. (1992). 'Extreme events in a changing climate: Variability is more important than averages'. *Climatic Change* **21**, 289–302.

Kim, H. J. and Siegmund, D. (1989). 'Likelihood ratio test for a change-point in simple linear regression'. *Biometrika* **76**, 409–423.

Leadbetter, M. R., Lindgren, G. and Rootzen, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. New York: Springer.

MacNeill, I. B., Tang, S. M. and Jandhyala, V. K. (1991). 'A search for the source of the Nile's change-points'. *Environmetrics* **2**, 341–375.

Mearns, L. O., Katz, R. W. and Schneider, S. H. (1984). 'Extreme high temperature events: Changes in their probabilities with changes in mean temperature'. *Journal of Climate and Applied Meteorology* **23**, 1601–1613.

Miller, K. A. and Glantz, M. H. (1988). 'Climate and economic competitiveness: Florida freezes and the global citrus processing industry'. *Climate Change* **12**, 135–164.

Potter, K. W. (1981). 'Illustrations of a new test for detecting a shift in mean precipitation series'. *Monthly Weather Review* **109**, 2040–2045.

Rukhin, A. L. (1994). 'Asymptotic minimaxity in the change-point problem'. In *Change-Point Problems. IMS Lecture Notes* **23**, 284–291.

de Rijk, W. G., Test, J. A., Penn, R. W. and Marsh, J. (1990). 'Applications of the Weibull method to statistical analysis of strength parameters of dental materials'. In *Progress in Biomedical Polymers.*, eds. C. G. Gebelein and R. L. Dunn. Plenum Press, 141–147.

Siegmund, D. (1988). 'Confidence tests in change-point problems'. *International Statistical Review* **56**, 31–48.

Weibull, W. (1939). 'A statistical theory of strength of materials'. *Proceedings of Royal Academy in Engineering and Sciences* **151**, 15.

Worsley, K. J. (1986). 'Confidence regions and tests for a change-point in a sequence of exponential family random variables'. *Biometrika* **73**, 91–104.

Yao, Y. C. and Davis, R. A. (1986). 'The asymptotic behavior of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates'. *Sankhya Ser. A* **48**, 339–353.

